1. Agyen, Vida Afarebea, Samuel Kobina Annim, and Emmanuel Ekow Asmah. 2024. “Neighbourhood Mothers’ Education and Its Differential Impact on Stunting: Evidence from 30 Sub-Saharan African Countries.” *Social Science & Medicine* 340:116462. doi: 10.1016/j.socscimed.2023.116462.

The authors use multilevel logistic regressions for this paper. They also utilize spatial analysis. They use the Demographic and Health Survey (DHS) dataset which has demographic and health indicators from 1990 to 2018.

The unit of analysis is individual-based, and each data point represents a child.

The authors use the removal method to handle missing data.

I find this helpful because it uses the same dataset that I am interested. The combination of statistical and spatial analyses is also worth noting.

<https://dhsprogram.com/data/available-datasets.cfm>

1. Adedokun, Sulaimon T., and Sanni Yaya. 2020. “Correlates of Antenatal Care Utilization among Women of Reproductive Age in Sub-Saharan Africa: Evidence from Multinomial Analysis of Demographic and Health Surveys (2010–2018) from 31 Countries.” *Archives of Public Health* 78(1):134. doi: 10.1186/s13690-020-00516-w.

The authors use multinomial logistic regression for this study, and they also used the Demographic and Health Survey (DHS) dataset from 31 countries.

The unit of analysis is individual based, while each data point represents a woman of reproductive age.

The authors did not state how missing data were handled.

This article broadly examines health and inequalities in Sub-Sahara Africa which is closely relevant to my research interests.

<https://dhsprogram.com/data/available-datasets.cfm>

1. Abedi, Vida, Oluwaseyi Olulana, Venkatesh Avula, Durgesh Chaudhary, Ayesha Khan, Shima Shahjouei, Jiang Li, and Ramin Zand. 2021. “Racial, Economic, and Health Inequality and COVID-19 Infection in the United States.” *Journal of Racial and Ethnic Health Disparities* 8(3):732–42. doi: 10.1007/s40615-020-00833-

The authors use bivariate, partial correlation, and linear regression while using data from multiple sources.

County is the unit of analysis in this study

The authors use the removal method to handle missing data. “Values not within three inter-quartiles were removed as part of the data pre-processing”

I selected this article because I am interested to know on how one can combine datasets from different sources for statistical analysis for a study

[https://data.census.gov/cedsci/table?q= Computer&tid=ACSST1Y2018.S2801&t=Telephone](https://data.census.gov/cedsci/table?q=%20Computer&tid=ACSST1Y2018.S2801&t=Telephone,)

<https://www.governing.com/archive/state-minority-population-data-estimates.html>

<https://worldpopulationreview.com/states>

<https://www.google.com/covid19/mobility/>